SF2A 2019
P. Di Matteo, O. Creevey, A. Crida, G. Kordopatis, J. Malzac, J.-B. Marquette, M. N'Diaye, O. Venot (eds)

THE ACTION SPÉCIFIQUE OBSERVATOIRES VIRTUELS FRANCE (VIRTUAL OBSERVATORY FRANCE SPECIFIC ACTION) IN THE OPEN SCIENCE CONTEXT

F. Genova[1]

Abstract. Astronomy has been, and continues to be, a pioneer of Open Science. It has established and maintains a disciplinary data sharing framework, the astronomical Virtual Observatory, which enables astronomers to discover the data useful for their research, to access them, and to use them with interoperable tools. Thanks to the work of the data producers and of the VO developers, astronomical data are FAIR (Findable, Accessible, Interoperable, Reusable). The *Action Spécifique Observatoires Virtuels France* (France Virtual Observatory Specific Action - ASOV) was created in 2004 by the CNRS-INSU, with CNES support, as the French chapter of the International Virtual Observatory Alliance (IVOA), which defines the Virtual Observatory standards at the international level, and a co-ordination structure at the national level. The ASOV supports French participation to the IVOA and technical exchanges between the French astronomical data and service centres. Its role and impact are described in the Open Science national and international context.

Keywords: Virtual Observatory, Open Science

1 Introduction

This paper was the contribution of the *Action Spécifique Observatoires Virtuels France* (France Virtual Observatory Specific Action - ASOV) to the 2019 meeting of the French Astronomical Society SF2A, which was held in Nice 14-17 May 2019. It describes the current Open Science context at the international and European levels, and the situation in France, with in particular the publication of the National Strategic Plan for Open Science in 2018 (Section 1). Section 2 addresses why astronomy is generally considered as one of the pioneers of Open Science. Section 3 describes how the ASOV fits into this landscape, and its impact at the national and international levels.

2 The Open Science context at the international, European and national level

Open Science has already had a long history at the political level. OCDE was mandated in 2004 to work on access to data from publicly funded research. They produced in 2007 their *Principles and Guidelines for Access to Research Data from Public Funding* (OECD 2007). The G7/G8 Ministries in charge of research published strong statements in 2013 on the transformation of the way to make research following the Open Data paradigm*. They continued since then to work together on the subject and to publish statements, with in particular the mention of the FAIR principles in 2017†. The FAIR Principles, which state that data should be Findable, Accessible, Interoperable and Reusable, were published in 2016 (Wilkinson et al. 2016). Their mention in an international ministerial statement in 2017 demonstrates how fast they became known and endorsed world-wide. In Europe, Carlos Moedas, Commissioner for Research, Science and Innovation, promoted in 2015 *Open Science, Open Innovation, Open to the World* as a vision for Europe (Directorate-General for Research and Innovation 2016).

[1] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 11 rue de l'Université, F-67000 Strasbourg, France
*`https://www.gov.uk/government/news/g8-science-ministers-statement`, Section 3 Open Scientific Research Data
†`http://www.g7italy.it/en/science-ministerial-meeting/`

In France, the Law for a Digital Republic, published on 8 October 2016, provides a legal framework to favour opening and circulation of data and knowledge. Another top-down incentive has been the questions about data policies and management among those which should be answered by Research Infrastructures candidates to the National Research Infrastructure Roadmap since its 2016 update (the same has been true at the European level for the ESFRI Roadmap questionnaire, also since its 2016 update). Several data infrastructures, such as the *Centre de Données astronomiques de Strasbourg* (Strasbourg astronomical Data Centre CDS), and the National Humanities Data Infrastructure *Huma-Num* (previously ADONIS and CORPUS), have been included in the National Roadmap since its inception in 2008, with the more recent addition of the *Earth System Research Infrastructure*, which is now called *Data Terra*. Recenly, the Ministry of Higher Education, Research and Innovation (MESRI) installed the *Comité pour la Science Ouverte* (Open Science Committee) in 2018, with 4 Colleges, *Publications*, *Research Data*, *Europe and international*, and *Skills and training*, with "Project Groups" on *Evaluation*, *Open and open Source software*, *Observatory of informational practices*, and *Building bibliodiversity*.

The MESRI published in July 2018 the National Strategic Plan for Open Science[‡], with three commitments: generalise open access to publications, structure research data and make it available through open access, and be part of a sustainable European and international Open Science dynamics. Among the measures concerning data, the following particularly resonnate with what we do in astronomy:

- Make open access dissemination mandatory for research data resulting from government-funded projects,
- Create the conditions for and promote the adoption of an Open Data policy for articles published by researchers,
- Among the structuring measures, develop subject-based and discipline-specific data repositories.

Two other structuring measures have to be noted because they impact the research and data provider communities:

- Generalise the implementation of data management plans in calls for research projects,
- Implement a certification process for data infrastructures.

Among recent progresses, one should also have in mind the rapid take-up of the FAIR Guiding Principles for Scientific Data Management. Since their publication in 2016, the principles have become the "normal" way to refer to the concepts which underline data sharing. An Expert Group set up by the European Commission published in 2018 a reference report and action plan *Turning FAIR into Reality* (Hodson et al. 2018). The rapid emergence and growth of the Research Data Alliance[§], an international organisation created in 2013 which aims at building technical and sociological bridges to facilitate the open sharing and reuse of data, is another sign of the world-wide interest for open science. At a more technical level, persistent identifiers such as DOIs (Digital Object Identifiers) and ORCID for identifying people facilitate the publication, attribution and citation of data.

3 Astronomy and Open Science

Open Science concerns all research fields, but they are not all at the same evolution stage. The cross-disciplinary usage of data is often put forward in the open science context, but it requires lots of work at the disciplinary level, since one needs to know the data to make it FAIR. Also, as demonstrated by what happened in astronomy, science data sharing enables a change in paradigm in the way science is done, first at the disciplinary level, by allowing one to discover data, and to access, reuse and combine them.

Open Science has somehow a "long" history over more than a decade, as explained, but astronomy begun to practice it long before. The *Centre de Données astronomiques de Strasbourg* was created in 1972, with the mission to "collect useful information concerning astronomical objects that is available in computerized

[‡]French version: `http://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf`, English version `http://cache.media.enseignementsup-recherche.gouv.fr/file/Recherche/50/1/SO_A4_2018_EN_01_leger_982501.pdf`

[§]`https://rd-alliance.org`

form; upgrade these data by critical evaluations and comparisons; distribute the results to the astronomical community; conduct research, using these data", that it continues to fulfil nowadays. The database of the *International Ultraviolet Explorer* IUE was also heavily used in early times: the satellite was in operation from 1978 to 1996, and the data in the database had been downloaded in average 5 times in 1994 (Wamsteker & Griffin 1995). These early successes were only the beginning, because data reuse is at the core of astronomical research, allowing combination of data from different instruments to understand the physical phenomena at work in the objects - multi-wavelength/multi-messenger astronomy - and studies of temporal variations.

Another strong asset of astronomy in the Open Science context is its capacity to develop and maintain standards through co-operation at the international level, like we do to develop facilities and instruments in international collaborations. The FITS format was published in 1981 (Wells Greisen & Harten), and the paper cites discussions at a "Workshop on Standards for Image Pattern Recognition" held in 1977. The standard is maintained under the aegis of the International Astronomical Union. FITS integrates data and metadata providing information about the observation, enabling data reuse. It allows the sharing of telescopic observations and the development and sharing of common tools to use the data. Although FITS is not perfect and often criticized, all the developments which came next have been made possible by the fact that astronomy shares a common data format.

The key next step for the "FAIRisation" of astronomical data has been the development of the Astronomical Virtual Observatory (VO), the disciplinary framework of interoperability standards and tools. The activity began around 2000, with the first funded projects in Europe, France, UK and USA starting in 2001. The interoperability standards are developed and maintained by the International Virtual Observatory Alliance[¶] (IVOA), which was created in 2002. The IVOA standards enable interoperability of data and of tools - one can for instance navigate seamlessly from images to tables to spectra when using VO-enabled applications and data services. This is an essential element for success: Genova et al. (2017), who compare interoperability frameworks established by different disciplines to identify commonalities and differences, states that "it is essential, for community uptake of data sharing, that data producers are enabled to share their data as well as users enabled to use the shared data."

Thanks to the data providers and VO developers, most astronomical data are open and FAIR, and they were so before the FAIR principles were defined. Data from most observatories are openly available, often after an embargo ("proprietary") period during which they are reserved to the team which proposed the observation. Data producers use FITS, which makes data Reusable, and provide their data in the VO. The VO enables astronomers to Find, Access and Interoperate data. As explained, applications are also interoperable. Other formats proposed as alternatives to FITS will have to reach the same maturity level to preserve existing capabilities.

4 The Action Spécifique Observatoires Virtuels France

The CDS had been involved in the VO development from the beginning, leading the first international *Interoperability Working Group* (2001-2002) under the aegis of the OPTICON network. This Working Group developed the first VO standard, VOTable (a standard to represent a table), which was finalised in April 2002. The IVOA was created in May 2002, and it took up the task of developing, agreeing on and maintaining interoperability standards.

The first INSU/Astronomy-Astrophysics (INSU/AA) strategic planning exercise (*Exercice de Prospective CNRS/INSU Astronomie Astrophysique*) organised after the creation of the IVOA was held in 2003. The importance of the Virtual Observatory was already well understood, which led to two recommendations:

- Create a coordination structure at the national level
- Begin to study the inclusion of modelling data in the VO

As a result, the *Action Spécifique Observatoires Virtuels France* ASOV was created by CNRS-INSU in 2004, with support from CNES. It has been regularly evaluated and renewed since then. With respect to the second recommendation, an IVOA Interest Group was set up in 2004 to work on requirements to include modelling data in the IVOA. The first assessment of this inclusion was taken up in the framework of the Euro-VO Data Centre Alliance (Euro-VO DCA) project. The project (project RI031675), coordinated by CDS on behalf of

[¶] `http://www.ivoa.net`

CNRS, was funded by the European Commission for 28 months from 2006 to 2008. French teams, led by Frank Le Petit and Hervé Wozniak, played an essential role in the definition of IVOA theory standards during and after EuroVO-DCA.

The ASOV has kept the same structure since its inception: the Scientific Council is composed of members designated by the Programmes and the other Actions Spécifiques, plus a handful of specialists. The Scientific Council chairperson represents France at the IVOA Executive Board[‖]. The ASOV role is to coordinate VO activities at the national level, and to disseminate knowledge and good practices about VO standards and tools. Following a recommendation of the 2014 Astronomy-Astrophysics strategic planning exercise, the ASOV mandate was expanded to include the coordination of technical exchanges between the teams which are engaged in data management, in particular for the ANO5 services[**], when it was renewed in 2016. This additional mandate was making explicit a function that ASOV already fulfilled.

The ASOV funds travel. It organises an annual call for proposals to support travels to IVOA meetings and to similar meetings of other sub-disciplines, and collaboration, thematic and regional meetings. The travels of the people who have responsibilities in the IVOA, with also travels from representatives of the main laboratories involved, are covered as far as possible. The ASOV also organises an annual meeting, which has been complemented since 2016 by a *Semi-Hack-a-Thon* for technical exchanges.

ASOV has been having a structuring role at the national level, and enabled the French community to be very active and visible in the IVOA. The CDS has been the ASOV starting point, and remains a cornerstone. But the VO community is now country-wide, as demonstrated by the fact that the six-monthly IVOA meeting was held at Paris Observatory in May 2019. Staff working in Besançon, Bordeaux, Grenoble, Marseille, Montpellier, Nice, Paris, Paris-Sud, Strasbourg and Toulouse participated in the ASOV 2019 meeting. There are currently a number of data services in France, many of them organised in Regional Expertise Centres, some of them labelled as ANO5 Observation Services. The ANO5 evaluation committee requires that data dissemination by the services is VO-enabled. Thanks to ASOV, knowledge of how to implement the VO framework is now widespread in the services and Regional Expertise Centres. French data services implement the VO standards and share good practices. The ASOV also encouraged regional exchanges when appropriate, in particular in the South-West (Bordeaux, Montpellier, Toulouse), as a precursor of the Regional Expertise Centre which was labelled in 2013.

The strong impact of the French community in the IVOA can be quantified: in May 2019, there were 45 IVOA standards. 27 of them have at least one author from a French laboratory/observatory (CDS/Observatoire Astronomique de Strasbourg, Paris Observatory, Grenoble, Montpellier and Toulouse), among which 17 have at least one editor from a French laboratory/observatory. French teams also develop key VO tools, in particular Aladin and Cassis, and Paris Observatory provides tools to validate standards, which play an essential role in the VO framework.

One important aspect of ASOV is that it covers all the sub-disciplines of the Section 17 of the *Comité National de la Recherche Scientifique*, namely astronomy, the study of the planets, the Sun and the heliosphere, space plasma physics, the astronomy facet of astroparticle physics, and atomic and molecular physics of astronomical interest. This enabled early dissemination of the astronomical VO concepts and tools, and the French community has been leading, or very active in, the development of disciplinary interoperability layers in European projects and at the international level in the nearby disciplines. One can cite the ASTERICS Cluster (2015-2019) with its *Data Access, Discovery and Interoperability* Work Package (Genova et al. 2019), which includes CTA, KM3Net, EGO-VIRGO, and was extended to EST at the end of the project, and its successor ESCAPE (2019-2023) with its *Connecting ESFRI projects to EOSC through VO framework* work package, which has a task on FAIRisation; Europlanet and VESPA (Virtual European Solar and Planetary Access) (Erard et al. 2018); and VAMDC, the Virtual Atomic and Molecular Data Centre (Dubernet et al. 2016). The other subdisciplines reuse or customize the standards developed for the astronomical Virtual Observatory for their own needs.

5 Conclusions

Astronomy is cited as an example in the context of Open Science and FAIR, for instance in Hodson et al. (2018). It created an international framework for FAIR data sharing, which is widely used by the community

[‖]The IVOA is composed of VO initiatives from countries, including France, plus the European Virtual Observatory collaboration and the inter-governmental organization ESA.

[**]ANO5 (ANO for *Action National d'Observation*) covers the centres in charge of data processing, archival and dissemination.

in its daily research work, and fully open: anyone can declare a service in the Virtual Observatory, or develop and share a tool to access VO-enabled data. Many current large projects provide their data in the Virtual Observatory. For instance, Gaia data was only made available in the VO by ESA, CDS, and other data centres which participate in the project data dissemination. The data from all ESA missions are VO-enabled, and VO standards are implemented in the new ESO programmatic interface (Romaniello et al. 2018). LSST staff participate in the IVOA, and contact with CTA, ELT, EST, Km3Net, SKA is well establisehd, in particular through ASTERICS and ESCAPE. Several space and ground-based observatories implement VO building blocks in their data pipelines. And in any case, all the data made available at CDS, including VizieR collection of catalogues and other data attached to publications and Aladin reference image archive, are available in the VO.

The ASOV enabled the creation of a national community of people engaged in the development of data services, and supported the community to influence and actively participate in the VO development. This was made possible because the national strategic planning exercises understood early the potential value of the Virtual Observatory and renewed their support, and the funding agencies, CNRS-INSU/AA and CNES, have been providing funds on the long term. Data is one of the research infrastructures of astronomy (Genova 2018), and like all research infrastructures, they require sustainable support.

The ASOV is grateful to INSU/AA and CNES, and to its community, for their continuous support along the years.

An earlier description of astronomy and the VO in the Open Science context was presented at the *Library and Information Services in Astronomy* (LISA VIII), held in Strasbourg 6-9 June 2017 (Genova 2018).

The author warmly thanks Franck Le Petit for his careful reading of the manuscript and his very relevant suggestions.

References

OECD 2007, Principles and Guidelines for Access to Research Data from Public Funding, OECD Publications, Paris

Wilkinson, M.D., Dumontier, M., IJsbrand, J.A., et al. 2016, Scientific Data, 3, 16001

Directorate-General for Research and Innovation (European Commission) 2016, Open innovation, open science, open to the world, Luxembourg: Publications Office of the European Union

Hodson, S., Jones, S., Collins, S., et al. 2018, Turning FAIR into reality, Luxembourg: Publications Office of the European Union

Wamsteker, W., & Griffin, R.E.M. 1995, Ap&SS, 228, 385

Wells, D.C., Greisen, E.W., & Harten, R.H., A&AS, 44, 363

Genova, F., Arviset, A., Almas, B.M., et al. 2017, Data Sci. J., 16, 16

Genova, F., Allen, M.G., Boisson, C, et al. 2019, PoS, à paraître.

Erard, S, Cecconi, B., Le Sidaner, P., et al. 2018, P&SS, 150, 65

Dubernet, M.L, Anthony, B.K., Ba, Y.A., et al. 2016, J.Phys.B, 49, 7/7/074003

Romaniello, M., Zampieri, S., Delmotte, N., et al., Proc. SPIE, 10704, 1070416

Genova, F. 2018, EPJ Web Conf., 186, 01001